

REMARKS

Claims 1-12 and 14-20 are pending in the application.

Claims 1-12 and 14-20 have been rejected.

Claims 1, 2, 6, 9-12 and 14 have been amended, as set forth herein.

New Claims 21-24 have been added.

I. **REJECTION UNDER 35 U.S.C. § 112**

Claims 1-5, 9-12, 14-15 and 18-20 were rejected under 35 U.S.C. § 112, first paragraph, as failing to comply with the written description requirement. This rejection is respectfully traversed.

Applicant has amended the Claims.

Accordingly, the Applicant respectfully requests withdrawal of the § 112 rejection of Claims 1-5, 9-12, 14-15 and 18-20.

II. **REJECTION UNDER 35 U.S.C. § 102**

Claim 6 was rejected under 35 U.S.C. § 102(b) as being anticipated by Sharman (US 5,774,854). The rejection is respectfully traversed.

A cited prior art reference anticipates the claimed invention under 35 U.S.C. § 102 only if every element of a claimed invention is identically shown in that single reference, arranged as they are in the claims. MPEP § 2131; *In re Bond*, 910 F.2d 831, 832, 15 U.S.P.Q.2d 1566,

1567 (Fed. Cir. 1990). Anticipation is only shown where each and every limitation of the claimed invention is found in a single cited prior art reference. MPEP § 2131; *In re Donohue*, 766 F.2d 531, 534, 226 U.S.P.Q. 619, 621 (Fed. Cir. 1985).

In general terms, Sharman recites breaking down words into syllables (dictionary look-up of the word to find the syllables), and then, breaking down the syllabified word into constituent phonemes (once again, a dictionary-lookup to find the phonemes). After the constituent phonemes have been determined, the acoustic processor determines diphones from the constituent phonemes, and looks at a diphone library (with prerecorded speech samples associated to the diphones) to determine the corresponding speech samples which are then concatenated. See, generally, Col. 5, line 18-40; Col. 6, lines 22-38.

Applicant's Claim 6 recites a method of pre-processing a text file. The received text file is parsed into textual units, "where each said parsed textual unit is one of a word, a prefix and a suffix." For each one of the parsed textual units, if one of the parsed textual units corresponds to a stored textual unit in a vocabulary of textual units wherein the vocabulary of textual units comprises words, prefixes and suffixes each having a pre-recorded speech sample associated therewith, then the stored textual unit is added to a list.

Sharman simply utilizes a dictionary look-up to determine the syllables of a word, then once the syllables are found, another dictionary look-up is used to determine the constituent phonemes, and then apparently phonemes are added to the list (which are input to the acoustic processor). In distinct contrast, only the parsed textual units (in the form of words, suffixes and

prefixes) are added to a list when the parsed textual units match a textual unit in a vocabulary of textual units, and where there is a pre-recorded speech sample associated with the textual unit. Sharman apparently recites when a textual unit (word) is matched, syllables are determined, then, when the syllables are matched, phonemes are determined and added to a list. Thus, Sharman's linguistic processor outputs "an annotated lists of phonemes" to the acoustic processor. Col. 6, lines 18-21. Applicant's invention (Claim 6) adds the stored textual units (in the form of words, prefixes and suffixes) to the list, and does not reduce to smaller sized units -- as is done by Sharman.

Moreover, Applicant's vocabulary of textual units comprises words, prefixes and suffixes each having a pre-recorded speech sample associated therewith. Even assuming that Sharman may disclose the parsed textual units as claimed by Applicant (which Applicant believes Sharman does not), Sharman fails to disclose textual units of words, prefixes and suffixes with each one having a pre-recorded speech sample associated therewith.

Therefore, Sharman fails disclose each and every element/feature of Applicant's Claim 6. Accordingly, the Applicant respectfully requests the Examiner withdraw the § 102(b) rejection of Claim 6.

III. REJECTION UNDER 35 U.S.C. § 103

Claims 1-5, 9-12 and 18 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Sharman (US 5,744,854) in view of Hata, et al. (US 5,878,393). Claims 7 and 16 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Sharman (US 5,744,854) in view of page 298 of Microsoft Press, Computer Dictionary (hereinafter “R1”). Claims 8 and 17 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Sharman (US 5,744,854) in view of R1 and further in view of O’Donnell (“Programming For The World - A Guide To Internationalization”, ISBN 0-13-722190-8). Claims 14-15 and 19-20 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Sharman (US 5,744,854) in view of Hata, et al. (US 5,878,393) and further in view of Malsheen, et al (US 4,979,216). The rejection is respectfully traversed.

In *ex parte* examination of patent applications, the Patent Office bears the burden of establishing a *prima facie* case of obviousness. MPEP § 2142; *In re Fritch*, 972 F.2d 1260, 1262, 23 U.S.P.Q.2d 1780, 1783 (Fed. Cir. 1992). The initial burden of establishing a *prima facie* basis to deny patentability to a claimed invention is always upon the Patent Office. MPEP § 2142; *In re Oetiker*, 977 F.2d 1443, 1445, 24 U.S.P.Q.2d 1443, 1444 (Fed. Cir. 1992); *In re Piasecki*, 745 F.2d 1468, 1472, 223 U.S.P.Q. 785, 788 (Fed. Cir. 1984). Only when a *prima facie* case of obviousness is established does the burden shift to the applicant to produce evidence of nonobviousness. MPEP § 2142; *In re Oetiker*, 977 F.2d 1443, 1445, 24 U.S.P.Q.2d 1443, 1444 (Fed. Cir. 1992); *In re Rijckaert*, 9 F.3d 1531, 1532, 28 U.S.P.Q.2d 1955, 1956 (Fed. Cir. 1993).

If the Patent Office does not produce a *prima facie* case of unpatentability, then without more the applicant is entitled to grant of a patent. *In re Oetiker*, 977 F.2d 1443, 1445, 24 U.S.P.Q.2d 1443, 1444 (Fed. Cir. 1992); *In re Grabiak*, 769 F.2d 729, 733, 226 U.S.P.Q. 870, 873 (Fed. Cir. 1985).

A *prima facie* case of obviousness is established when the teachings of the prior art itself suggest the claimed subject matter to a person of ordinary skill in the art. *In re Bell*, 991 F.2d 781, 783, 26 U.S.P.Q.2d 1529, 1531 (Fed. Cir. 1993). To establish a *prima facie* case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations. The teaching or suggestion to make the claimed invention and the reasonable expectation of success must both be found in the prior art, and not based on applicant's disclosure. MPEP § 2142.

a. Claims 1-5, 9-12 and 18

As described in the previous section, Sharman parses the text file down to constituent phonemes, and then matches the phonemes to diphones in a diphone library to find associated speech samples. The speech samples of Sharman are prerecorded speech segments associated with diphones. Hata, on the other hand, uses whole words or phonemes without reference to prefixes and suffixes (or root words).

Parsing the text in the manner described by Applicant, and which is different than that described in Hata (whole words), reduces the number of textual units and associated speech samples that the system needs. As described in the specification (page 6), a pre-defined vocabulary of about 2300 words (in one example) were found experimentally relevant in an analysis of a large volume of email messages. By utilizing prefixes and suffixes, the number of combinations is increased, while still utilizing a manageable number root words, prefixes and suffixes. Moreover, parsing each textual input (text file and/or word) into phonemes and diphones, such as done in Sharman (and contemplated by Hata), significantly increases processing requirements to perform the text to speech conversion process. Applicant's invention utilizes pre-recorded speech samples associated with textual units larger than phonemes (unlike Sharman and Hata), and further utilizes prefixes and suffixes (unlike Hata).

In addition, neither Sharman or Hata disclose, teach or suggest locating an associated speech sample in memory, where the memory comprises a vocabulary of words, prefixes and suffixes and a plurality of speech samples with each speech sample corresponding to a one of said words, prefixes and suffixes (See, independent Claims 1, 9-12 and 14). Also, neither Sharman or Hata disclose, teach or suggest that when no word, prefix or suffix (each having an associated speech sample) can be found in the vocabulary, then marking/indicating the textual unit, and sending the "out of vocabulary" textual unit to a secondary text-to-speech (TTS) engine to generate a speech sample for that particular textual unit (See, dependent Claims 2, 3 and 18). The passage cited by the Office Action (Col. 5, line 24) referring to a "back-up

mechanism” appears limited to a mechanism that can determine the syllables of a word for a particular word that is not found in the dictionary. The passage cited by the Office Action (Col. 5, line 34) referring to “general rules for words not in the dictionary” appears limited to a mechanism that can determine the phonemes of a word that is not found in the dictionary.¹ Sharman’s description fails to describe that when the word is not found in the vocabulary (a vocabulary that has speech samples associated with each entry in the vocabulary), the “out of vocabulary” word is forwarded to a secondary TTS engine. Applicant suspects that Sharman simply desires to determine the syllables and phonemes (by some undescribed means or method) and then continue with generating the speech samples using the diphone library look-up (and phonemes). Sharman does not disclose sending the “out of vocabulary” word to a secondary TTS engine that generates a speech sample (for concatenation) - as Sharman breaks down the text into phonemes, and the phonemes are used to generate the speech sample. Thus, according to the Office Action interpretation, Sharman’s main device and the “back-up mechanism” would have duplicate functions. Applicant respectfully submits that this interpretation is incorrect and not supported by Sharman.

b. Claims 7, 8, 16 and 17

For the reasons set forth above establishing that independent Claim 6 is not anticipated by Sharman, and for the reasons set forth above with respect to Applicant’s secondary TTS engine, each of the proposed combinations of Sharman-R1 and Sharman-O’Donnell fails to

¹ None of the cited passages describe any specific method or device that would perform the function of determining the syllables or phonemes of a word not found in the dictionary.

disclose, teach or suggest all the elements/features of Applicant's invention (as claimed) and there is no motivation or suggestion to combine such references.

c. Claims 7, 8, 16 and 17

For the reasons set forth above in response to the Office Action's Section 103 rejection of independent Claims 1 and 9-12, as noted, the combination of Sharman-Hata fails to disclose teach or suggest all the elements/features of Applicant's invention (as claimed) and there is no motivation or suggestion to combine such references. Moreover, Malsheen fails to cure the deficiencies in Sharman and Hata.

Accordingly, the Applicant respectfully requests withdrawal of the § 103(a) rejection of Claims 1-5, 7-12 and 14-20.

IV. NEW CLAIMS 21-24

Applicant has added new Claim 21-24.

Independent Claim 21 provides a method of converting text to speech, and includes the step of determining whether the textual unit matches a stored textual unit in a vocabulary of textual units, and if so, retrieving from memory a speech sample associated with the stored textual unit, and if not, sending the textual unit to a secondary text to speech engine and receiving a speech sample converted from the textual unit from the secondary text to speech engine. None of the cited references disclose, teach or suggest this element/feature.

Independent Claim 23 recites memory (having an array of stored textual units each having an associated stored speech sample), a text pre-processor, a concatenation engine (operable for receiving one of the parsed textual units and retrieving a speech sample corresponding to the parsed textual unit and outputting the retrieved speech sample), and a secondary text to speech (TTS) engine. The secondary TTS engine is operable for receiving from the concatenation engine a parsed textual unit not found in the array of stored textual units, for converting the received parsed textual unit into a converted speech sample, and sending the converted speech sample to the concatenation engine, and further reciting that the concatenation engine retrieves the associated stored speech sample when the parsed textual unit is found in the array of stored textual units.. None of the cited references disclose, teach or suggest this element/feature.

V. CONCLUSION

As a result of the foregoing, the Applicant asserts that the remaining Claims in the Application are in condition for allowance, and respectfully requests an early allowance of such Claims.

If any issues arise, or if the Examiner has any suggestions for expediting allowance of this Application, the Applicant respectfully invites the Examiner to contact the undersigned at the telephone number indicated below or at *rmccutcheon@davismunck.com*.

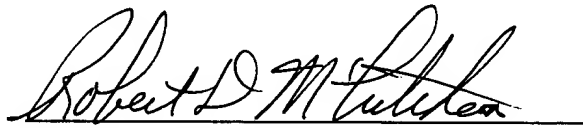
The Commissioner is hereby authorized to charge any additional fees connected with this communication or credit any overpayment to Davis Munck Deposit Account No. 50-0208.

Respectfully submitted,

DAVIS MUNCK, P.C.

Date: _____

9/7/2004



Robert D. McCutcheon
Registration No. 38,717

P.O. Drawer 800889
Dallas, Texas 75380
(972) 628-3632 (direct dial)
(972) 628-3600 (main number)
(972) 628-3616 (fax)
E-mail: *rmccutcheon@davismunck.com*